CORA Cork Open Research Archive
Cartlann Taighde Oscailte Chorcaí

Title	"What can I help you with?": infrequent users' experiences of intelligent personal assistants
Authors	Cowan, Benjamin R.;Pantidi, Nadia;Coyle, David;Morrissey, Kellie;Clarke, Peter;Al-Shehri, Sara;Earley, David;Bandeira, Natasha
Publication date	2017-09
Original Citation	Cowan, B. R., Pantidi, N., Coyle, D., Morrissey, K., Clarke, P., Al-Shehri, S., Earley, D. and Bandeira, N. (2017) '"What can I help you with?": infrequent users' experiences of intelligent personal assistants', MobileHCI '17: Proceedings of the 19th International Conference on Human-Computer Interaction with Mobile Devices and Services, Vienna, Austria, 4-7 September, Article 43 (12pp). doi: 10.1145/3098279.3098539
Type of publication	Conference item
Link to publisher's version	https://mobilehci.acm.org/2017/ - 10.1145/3098279.3098539
Rights	© 2017, Association for Computing Machinery. This is the author's version of the work. It is posted here for your personal use. Not for redistribution. The definitive Version of Record was published in MobileHCI '17: Proceedings of the 19th International Conference on Human-Computer Interaction with Mobile Devices and Services: https://dx.doi.org/10.1145/3098279.3098539
Download date	2023-05-05 10:09:53
Item downloaded from	http://hdl.handle.net/10468/8253

UCC
University College Cork, Ireland
Coláiste na hOllscoile Corcaigh

“What can I help you with?”: Infrequent users’ experiences of Intelligent Personal Assistants

1st Author Name
Affiliation
City, Country
e-mail address

2nd Author Name
Affiliation
City, Country
e-mail address

3rd Author Name
Affiliation
City, Country
e-mail address

ABSTRACT
Intelligent Personal Assistants (IPAs) are widely available on devices such as smartphones. However, most people do not use them regularly. Previous research has studied the experiences of frequent IPA users. Using qualitative methods we explore the experience of infrequent users: people who have tried IPAs, but choose not to use them regularly. Unsurprisingly infrequent users share some of the experiences of frequent users, e.g. frustration at limitations on fully hands-free interaction. Significant points of contrast and previously unidentified concerns also emerge. Cultural norms and social embarrassment take on added significance for infrequent users. Humanness of IPAs sparked comparisons with human assistants, juxtaposing their limitations. Most importantly, significant concerns emerged around privacy, monetization, data permanency and transparency. Drawing on these findings we discuss key challenges, including: designing for interruptability; reconsideration of the human metaphor; issues of trust and data ownership. Addressing these challenges may lead to more widespread IPA use.

Author Keywords
Intelligent personal assistants, speech interfaces, user experience, trust, privacy

ACM Classification Keywords
H.5.m. Information interfaces and presentation (e.g., HCI): Miscellaneous

INTRODUCTION
An Intelligent Personal Assistant (IPA) is “*an application that uses input such as the user’s voice… and contextual information to provide assistance by answering questions in natural language, making recommendations and performing actions*” [3, p.223]. Within the literature the term IPA is used interchangeably with terms such as Conversational Agents, Virtual Personal Assistants, Personal Digital Assistants, Voice-Enabled Assistants or Voice Activated Personal Assistants, to name a few. IPAs combine speech recognition, language understanding, dialogue management, language generation and speech synthesis to respond to user queries and requests. Voice enabled IPAs like Siri, Google Assistant, Microsoft Cortana and Amazon Alexa are widely available on smart phones, and increasingly in homes (e.g. Amazon Echo and Google Home) and cars (e.g. Google Assistant integration with Hyundai). The market for IPAs is predicted to reach $4.61 billion by the early 2020s [25].

The technical infrastructures that enable IPAs have advanced rapidly in recent years and have been the subject of extensive research (e.g. [9,12]). However, research focused on understanding the user experience of IPAs is more limited [1]. Unsurprisingly, from our perspective, this has had implications for the adoption and use of IPAs. For example, despite their widespread and widely promoted inclusion on mobile devices people tend to use IPAs rarely or not at all [42]. A recent survey showed that 98% of iPhone users had used Siri in the past. However only 30% used it regularly, with 70% using it rarely or only occasionally [16]. The trend is similar for other IPAs [16].

Previous work has investigated the use of IPAs by distinct user groups, including children [30] and older adults [41]. More recently Luger and Sellen [31] provided a comprehensive study of the experiences of people who use IPAs on a frequent basis (almost daily). Such studies are invaluable, but in isolation they cannot explain the experiences and barriers to use of all potential users. In this paper, we adopt a distinct approach, aimed at understanding the experience of people who use IPAs occasionally, but not frequently. Surveys suggest this reflects the most common pattern of use (70% of all users). For those wishing to extend the use of IPAs – including Google, Apple and Amazon – these infrequent users represent a key target group and one not previously studied in detail.

This paper is grounded in a series of focus groups where 20 participants, most of whom were infrequent users, discussed their experiences using an IPA on a smart phone. The analysis focused on users’ views, everyday practices and barriers to use and resulted in six core themes: 1) issues with supporting hands free interaction; 2) problems with performance with regards to user accent and speech recognition more widely; 3) problems around integration

Paste the appropriate copyright/license statement here. ACM now supports three different publication options:

- ACM copyright: ACM holds the copyright on the work. This is the historical approach.
- License: The author(s) retain copyright, but ACM receives an exclusive publication license.
- Open Access: The author(s) wish to pay for the work to be open access. The additional fee must be paid to ACM.

This text field is large enough to hold the appropriate release statement assuming it is single-spaced in Times New Roman 8-point font. Please do not change or modify the size of this text box.
Each submission will be assigned a DOI string to be included here.

with third party apps, platforms and systems; 4) social embarrassment being a barrier to using mobile IPAs in public; 5) the human-like nature of IPAs; and 6) issues of trust, data privacy, transparency and ownership. Our findings complement previous themes found with power users of IPAs but are the first to contribute towards understanding the experience of infrequent IPA users. They provide helpful insights by beginning to understand which issues are acceptable frustrations for some users, but key barriers to adoption for other potential users. The themes regarding concerns over trust, transparency and privacy have not been identified in previous studies and may be a key difference in the adoption choices of frequent versus infrequent IPA users. They thus point towards key unmet challenges in extending IPA use.

Understanding and addressing the experiences and barriers to adoption amongst infrequent users is critical to fully realising the potential of IPAs. If these barriers cannot be addressed, developers may be best served in accepting that IPAs will remain a relatively niche application and focus on maximising the experience of power users. Ultimately, we believe enhanced user experiences and widespread adoption is an achievable target. This paper aims to support this goal.

RELATED WORK

IPAs operate on the same fundamental principles as many other spoken dialogue systems. They firstly recognize the user's utterance (through speech recognition) and interpret the meaning of the words and phrases recognized (language understanding). The IPA then selects an appropriate action to perform based on the meaning of the phrase recognized, the current state of the dialogue and what has already been shared between the user and the system (through a dialogue manager). The content that the IPA will deliver to the user is then planned and a response is generated and turned into speech (through speech synthesis) and/or text output [24,29]. Some IPAs, like current versions of Siri and Google Assistant, also use contextual information that can be acquired from mobile devices to tailor options and information given in response to queries [32]. Interaction with IPAs tends to take on a question and answer format that is highly task oriented, rather than being social or conversational, aside from a number of pre-scripted jokes and humorous responses.

People's experiences of IPAs

Despite their widespread availability and improved technical capabilities, there has been little work on people's experiences of IPAs. However interest within the HCI community has been growing [11,31]. For example, Wulf et al. [41] investigated the potential of speech only interfaces (Siri with the screen occluded) for older users. The authors focused on common tasks like asking for the weather, creating notes, asking for directions, writing a text message, asking for an address and setting an alarm and found that participants were quite positive to the idea of using speech interfaces. While issues were identified in relation to the quality of the Internet connection, people forgetting to use the button to activate Siri and the quality of the speech recognition, participants were often surprised at how effective and fast interaction with the IPA was.

A second recent study explored children's use of Siri [30] and found that they predominantly ask Siri questions and requests that focus on getting to know or exploring the agent (e.g. asking it personal questions), getting specific information about a topic and using Siri to make a call or send a text, although the last activity was less common. Children also seemed keen to test the limits of Siri, by asking questions such as "where is mommy?" that were out of its capability. One of the major findings from this work was that Siri had limitations in recognizing children's speech (this is a significant technical challenge in the field of speech recognition - see [38,39]). When their speech was not recognized, children used strategies like speaking more loudly, making threats to the IPA and pausing after each word to try and get it to recognize their speech.

Studies such as [30,41] go some way to shedding light on the user experience of IPAs, but have tended to focus on specific user groups. A more comprehensive, qualitative study was recently presented by Luger and Sellen [31]. Through semi-structured interviews the study explored how a set of frequent users used IPAs, where they use them and the emotions elicited by their use. Their work found that these frequent users use them in hands /eyes busy situations (like driving or looking after children) and that the interaction with the agent was generally seen as a secondary task. These users also did not trust the system to do complex tasks like writing emails or calling someone, down to an apprehension that the system would not get the task done correctly. They also found that they used strategies like dropping colloquialisms or complex words, reducing the number of words used, altering enunciation, speaking more clearly or slowly and changing the accent used when interacting with IPAs. These echo the findings of previous work on language choices in human-computer dialogue [4,27,35]. Importantly, these users saw IPAs as an interface that had to be learned. The work also highlights people's prevalent perceptions of the agent as a non-competent interface and a lack of a clear mental model of the agent with these users being uncertain about what the system could do, how it worked and whether it changed over time with more interaction. The paper suggests that anthropomorphism also raises expectations of human abilities in such interactions. The user's mental model in interaction is a critical issue in speech interface interaction [3,6,13], heavily influencing our interaction behaviours [6,35].

Finally, the question of context of use has also been explored within the literature available on IPA use. Studies [33,34] have shown a clear need to consider the public nature of the context of use and the information being transmitted in such interactions. In a quantitative rating

study of imagined interaction scenarios, people rated themselves more highly likely to use an IPA in a private place, as well as when disclosing non-private information. People also rated IPAs as more acceptable to enter non-private information and to use it at home, with less of a difference in acceptability rating for private and non-private information in home than public contexts.

As presented above, studies that have focused on understanding people's experience and use of IPAs are limited. At the same time, IPAs are becoming more widespread in various everyday applications and contexts (e.g. home, car), but many people still do not use them. If IPAs are to become a more mainstream interaction, we need to understand the key reasons why people currently do not use them or use them infrequently. There is clearly a need for further work on people's experiences of IPAs so as to identify and address the current state of the IPA user experience. Our study aims to address this by presenting findings from a focus group study investigating how infrequent users experience and interact with IPAs. In the following sections, we detail our method, present the findings of our data analysis and discuss their implications.

METHOD

Participants

We recruited 20 participants (11 Female, 9 Male) from a university community (14 students; 6 non students). The majority (65%) of the participants were [*specific nationality*], with all participants being native or near native English speakers. Participants came from a variety of academic backgrounds and varying levels of technical proficiency, with most rating their technological proficiency as advanced (55%) or intermediate (40%). We focused our recruitment on Siri users specifically. This was for two reasons. Firstly, Siri is the most commonly used IPA, with 71% of participants in previous studies citing that they used Siri [31]. Focusing on one IPA also ensured that all participants in the focus groups had a shared device context to reflect on rather than reflecting on varied IPA interfaces. Similar to trends observed in consumer surveys on Siri use [16] the majority of our participants reported that they were not currently active users (55%). A number of users also stated that they used Siri less than once a month (25%) with only one of our participants using Siri more than once a day. Participants were each given a €10 honorarium for taking part.

Procedure & Analysis

Our study involved two stages. First, participants filled in an online questionnaire and completed six common tasks with Siri. The online questionnaire solicited demographic information (e.g. age, nationality, education, profession), their level of technical experience, how often they used Siri and what they tend to use Siri for. As part of the questionnaire, people were also given six tasks to complete, using Siri. For each task participants were asked to report any comments, issues or observations they had when performing the tasks. The full task list is shown in Table 1. These tasks were chosen based on common uses of Siri found in previous work [31] as well as those suggested in Siri user guides and Apple advertising. The tasks were intended to help prompt discussion amongst participants in the subsequent focus groups as well as potentially remind participants of Siri's functionality, as many of our participants were infrequent users.

	Task
1	Find out what the weather will be like in [*city name*] tomorrow
2	Get driving directions to [*city name*] city centre
3	Send a text to one of your contacts
4	Find out the number of calories in the last meal you ate
5	Set a reminder
6	Search for a recipe

Table 1: The tasks participants were asked to conduct with Siri before attending the focus group.

Participants then took part in semi-structured focus groups, which were used to elicit their views and everyday practices around using Siri. We chose to use focus groups and limited their size to 3-5 participants per group to capitalize on emergent dialogue between participants [28]. Each focus group lasted approximately 60 minutes. Focus groups were run until data saturation was reached. The focus groups took place over a period of four weeks. Participants were randomly assigned to the five groups. All participants participated in a focus group that took place on the day after they completed the online questionnaire and six tasks listed in Table 1.

Each focus group began with an icebreaker and continued with questions on the following topics: 1) *general perceptions of intelligent personal assistants, 2) reflections on the set of tasks conducted before the focus group, 3) their first impressions of Siri, 4) their interaction with Siri, 5) the context of use, 6) any issues with using Siri* and 7) *potential future scenarios*. The collected data was transcribed and Inductive Thematic Analysis [7] was used to analyse the data. Our data analysis followed the six steps used for inductive thematic analysis as described by Braun and Clarke [7]: i) Familiarisation with data (by repeated and active reading), ii) generation of initial codes, iii) sorting initial codes and forming themes, iv) reviewing themes, v) defining and naming themes, vi) identifying potential relations/hierarchies between the themes. The data analysis was conducted by multiple coders that worked independently. Four researchers each analysed the data from the focus groups and then one researcher, with extensive experience in qualitative analysis, analysed the data from all focus groups blindly (without having seen the analyses from the four others- blind double coding). These two sets of independent analyses were followed by two data sessions where two additional researchers joined the above

in scrutinizing the resulting analytic themes, making sure they represented accurately the patterns of meaning that occurred during the focus groups. All involved researchers have a background in HCI, three of them with a background in Psychology; from those, two have extensive experience (6+ years) in conducting qualitative work.

PRE-FOCUS GROUPS TASK

From the online questionnaire, on the topic of what our user group tend to use Siri for, the following tasks were reported: checking the weather, directions, sports results, setting reminders and alarms, turning data off, making calls or sending short messages when hands are busy, searching the web, opening apps or for fun (e.g. asking it stupid questions to get answers or getting it to tell a joke).

In relation to their experiences with Siri in the tasks, a number of our participants were impressed by Siri's performance, although it was not uncommon for users to report a number of tries being needed to complete the tasks. Issues were encountered where Siri did not give accurate information (e.g. they received weather forecasts for the wrong time period or got directions to the wrong destination). Some were frustrated when Siri asked for options to be selected on the screen. Both of these issues were seen to compromise the hands free nature of Siri. A number of observations were also made about Siri sometimes not speaking answers but bringing up visual information or search results, again impacting the hands free nature of interaction. This topic was also raised as a key issue in the focus groups. Siri also sometimes misrecognised what people had said, thus bringing up irrelevant results or setting up messages to be sent to the wrong contact. There was also frustration that apps that were not used by our participants, like Apple Maps or Bing, were used as defaults for related user queries. In the text message task, our participants also described how Siri seemed to work better for short than long messages. A number of participants made specific reference to its ability to understand their accent and recognising the words they spoke to the system when commenting on Siri's performance. Some also felt that they got a more natural response from Siri if they themselves were more natural in what they said to it. These issues were explored in more detail in the focus groups.

FOCUS GROUP FINDINGS

We find six key themes in our data that serve to highlight issues in the Siri user experience for infrequent users. These focus on 1) the hands free nature of Siri use and how this is compromised by the need for screen input by the user; 2) issues with speech recognition and strategies used when interacting with Siri; 3) poor integration with third party apps, platforms and systems; 4) social embarrassment as a key reason for Siri not being used in public; 5) the positive and negative impact of Siri's human like nature on the user experience; 6) concerns over data privacy, confidentiality, monetization, data permanency and user tracking. We detail each of the themes uncovered in our data in the following sections.

Hands-*not*-free interaction

Participants saw Siri as being most useful in situations where they are not able to have full use of their hands – e.g. when driving – and for simple tasks, such as dictating a short text message and setting reminders.

While using Siri in hands-free contexts was the most commonly mentioned, participants were quick to identify that in practice Siri is not fully hands-free. Many of them described daily occurrences where "hands free" included having to tap to unlock the phone or to choose from a list of names to send a text message.

"*When I was trying to send a message to my brother, it made me unlock my phone, which was annoying because I'm trying to be hands-free, I'm not trying to unlock my phone. And if I was just doing it while not looking at it, (…) if I was in a car or something, I wouldn't be looking at it so I wouldn't know that it hadn't sent the message. And then if I tried to edit the message, it just has me redo the whole thing, which I thought was stupid.*" [FG2, P2]

"It's like when I was sending a text, I was like "send John whatever text" and then it was like, "tap on the John you want to send it to," and then I'm like, "well, if I'm going to tap it I might as well just write the text." [FG3, P4]

Issues around interrupting the hands free nature of the interaction were reported by our participants both during the focus groups and before (when asked to complete the tasks with Siri) and were found to be prominent in obstructing potential frequent use.

User Accent and Speech Recognition

Siri's performance regarding recognising users' accents was, as expected, a prevalent topic of discussion. Participants talked about noticeable system improvements, but equally referred to examples of Siri misrecognising what they have said. The majority of participants noted that Siri's speech recognition performance has improved significantly since its very first release:

"She has, she got better, like, if I... my dad had an iphone 6, sorry 5 and I think that was when Siri was introduced first and he couldn't get it to work. You had to be really specific with what you wanted Siri to do. So now you can just say 'turn on the alarm' and it gives you a list of alarms, but before it would just go 'Woahh alarm? what is an alarm?'" [FG1, P4]

"I'd be the same in kind of, watching the progress of it, because I've had iPhones for years and to see how much it's starting to pick up on what I'm saying. I don't have a particularly strong accent, but the first one that came out, it hadn't the foggiest what I was saying." [FG3, P5]

That said, several speech recognition issues seemed to be persistent. Participants brought several examples of how Siri still does not accurately pick up what they say especially when they are in noisy environments. Many attributed these issues to Siri not dealing well with accents that are local or non-native:

"If you're in a bar or anywhere with any kind of degree of noise, city centre, bars, restaurants, it's just very hard for it to understand you (...) and it just kept saying "mail," and I'm like "meal," saying it clearly like, but it just kept going "mail." Like, you're getting a bad mark for that." [FG2, P1]

"I think it does struggle with some of our idioms and our accents. It just doesn't understand. So there's a few of the tasks that I had to repeat myself several times, or had to phrase things differently because it wasn't picking it up." [FG3, P2]

In the same context, participants reported having to think carefully about their phrasing before addressing Siri as pausing and single word editing were problematic:

"*I feel like you kind of have to take a few seconds to think about what you're going to say first. Because if you do it the normal way where you're talking to someone and you pause, then it'll just start doing whatever, instead of waiting for you to actually finish"* [FG3, P4]

"*Yeah, when I was trying to edit it, it was just one word that it got wrong. I said "patterns" and it said like "patters," and so I was just trying to fix one word. I was asking my brother to send me a list of patterns, and I couldn't get Siri to just edit one word. I had to say the whole message again and I was like, this is not really good.* [FG2, P2]

Participants also mentioned several strategies they use to prevent speech recognition issues and maximise Siri's performance such as phrasing things in a specific way, speaking slowly and hyperarticulating.

Integration with apps, platforms and systems

Participants across all focus groups discussed issues with respect to Siri's interoperability with third party apps, platforms and devices and expressed a strong preference that Siri becomes more customisable and better integrated.

A main source of frustration for many of the participants was Siri's default use of Apple apps – its *appleness* - and lack of integration when it comes to third party apps. Participants found this very limiting and commented on how they would be more likely to use Siri if there was the option to customise its use across other apps and platforms:

"Because right now Siri's limited to only Apple products, but I hate Bing, and I hate Apple Maps, and the only person that has iMessage is my brother, so it's sort of limited by its Appleness. (...) If she used Google Maps, I might be more likely to ask her for directions if I'm going somewhere but she doesn't, so I don't. (...) and to add Messenger interfunctionality, because I have like, six messaging applications on my phone. If I could say, send a telegram message to Marcus, send a WhatsApp message to my dad, send an iMessage to my brother, because everybody uses different platforms. If I could tell Siri to use a certain application to send a message I would probably use it more" [FG2, P2]

"It just opens Apple Maps, which is insane. Because it's terrible. I don't know why they didn't ask which maps would you prefer, or have a preference setting, like set a customization.(...) It's the new iOS that they brought it in, and it's the worst thing that they could have done. Well, you can activate Google by saying "please Google this for me" but I'm never going to think of that." [FG3, P3]

A few participants also wished Siri's functionality could be extended across other Apple products:

"Because I have all the Apple things, I have the MacBook, the iPad, I want Siri to be able to be like, interact with my MacBook, because I'm usually on my phone and my computer at the same time. So it would be great if Siri could like, I have an iCloud account and they're connected anyway, it would be great if I could be like "oh, open this thing on my computer." [FG3, P2]

The lack of integration and ability to customize to their regular app usage habits was a key barrier in the eyes of our participants. Integration across devices was also seen as a potential way to improve the overall experience.

Public Use and Social Embarrassment

Participants reported avoiding using Siri in public or outdoors spaces. While issues of recognition were related to this, the main reason for not using Siri in public was *social embarrassment*. Talking to Siri was discussed as different to talking to someone on the phone, it was more like "*talking to a wall"* or yourself and that it felt weird or inappropriate to use Siri in public spaces or in the presence of others, especially strangers.

"If you're talking to a person, it's not awkward. If you're on the phone, it's not awkward. But you feel like you're talking to the wall. So there is something awkward about putting yourself out there, because you're talking to something that doesn't really exist. And then it's strange to kind of, be talking to nobody. Because you are just talking to yourself. You're not giving instructions to a person, you're not having a conversation. It's normal to have a conversation with someone when it's quiet, but it's not normal to just speak aloud to yourself." [FG3, P1]

"*You know if you're using it in public you might be embarrassed or, I'd be kind of conscious of that now, like talking to a phone and it talking back to me. Especially on the bus now, I wouldn't use it"* [FG1, P3]

Participants' social embarrassment concerns further implied a strong *cultural element*; many mentioned that in an *[specific nationality]* context talking to your phone in

public is not as socially acceptable and would often be criticized by onlookers:

"*Why I don't use it? Because I'm [specific nationality] I think, I'm not going to be sitting on a bus saying ya know... I feel I would just be conscious about that."* [FG1, P4]

The potential for social embarrassment by using Siri in public seemed to be a major concern and an obstacle to our user group using this in public. There also seemed to be a cultural element to this, whereby this type of embarrassment was framed as being caused by cultural norms on phone use in public.

Human-Like Nature of Siri

While discussing their interaction with Siri, participants commented and debated extensively around human-like features of Siri such as personality and how these impact or create expectations around the interaction. Participants assigned different personalities to Siri, discussed those with respect to its voice and accent, considered conversing with versus commanding Siri and engaged in comparisons between Siri and humans as personal assistants.

Personality and Voice

Several participants assigned Siri human traits and discussed how having a personality makes Siri more user friendly:

"She's always really sassy, like when she turns on my iPad, and she's like "Oh, what can I help you with?" [FG2, P2]

"I think having a personality helps a lot in making it less than just, like you're barking commands into your phone" [FG1, P5]

Others disagreed and discussed extensively how what is understood as personality is a design feature hard coded by its creators to make Siri *appear human* more relatable and therefore more user friendly without always being successful in doing that:

"the tasks that I used it for, I didn't see a personality really in it." [FG5, P1]

"I think they try to imbue a slight sass into it, especially if you ask it questions about Apple, or love (...) It's a weird kind of hard-coded personality where it's very set questions where it has a personality, and then all the others ones it doesn't. So it chooses when to have one." [FG3, P4]

Siri's human-like voice was one feature that was generally agreed to make it more human-like and user friendly. Siri's voice was found to be less robotic and closer to that of a human person. Participants further picked up on how cultural nuances have been imbued in Siri's voice and personality; therefore, choosing an accent (e.g. Australian, British, Japanese) delivers a different user experience altogether:

"You know, it seemed to kind of engage with you, it wasn't such a robotic, it was a real kind of empathy, is the word, behind the voice as well." [FG2, P1]

Even the voice you choose, so if you're "Australian male," that's going to seem like a different experience, to like, the British woman is really posh. [FG3, P1]

"I switched her over to Japanese and she is a lot more serious and polite, and less likely to crack jokes. I think there is a personality for each of the different languages depending on what that language requires I guess (...) unlike the English version, the Japanese don't really do that, it's not really colloquial in the same way, it's very formal. It's very much an assistant rather than a personality on your phone, if that makes sense." [FG1, P5]

Intelligent Personal Assistants vs Human Assistants

Siri's human-like traits also sparked several comparisons between what Siri as a personal assistant can do or should be able to do versus a human personal assistant. Several participants commented that Siri is still very limited compared to humans:

"Like scheduling meetings. I think it's a long way from that, I think it's not good enough yet. And there's a lot of things that a person can do that Siri can't, like the nuances of scheduling and privatising, things like that. [FG5, P1]

"No, Siri's not a real person. She doesn't remember what I say to her, she probably logs it for data purposes but she's not going to be like, "do you remember that time you told me this thing?" [FG2, P2]

Others argued that Siri can replace a human personal assistant, but instead the design ambition should be making Siri better than humans:

"...maybe we should stop trying to make Siri like a person, because a person can only do so much. Make it better than a person, more helpful. Because if Siri's supposed to be a personal assistant, I would rather a personal assistant that can do way more than a human." [FG3, P3]

Participants also discussed Siri's limitations with respect to its conversational abilities. Currently users' interaction is more about giving commands rather than having a conversation as you would do with another person:

Because it's not natural language. It's commands, you're giving a command to it, and you have to make it as specific as possible. So I think it is quite noticeable that you're not talking to a real-life human being." [FG3, P2*]*

Following from this, several participants expressed a desire to be able to converse with Siri as this would be more fun and useful, while others were sceptical about the boundaries of such an interaction:

"I would like Siri to ask me back a question, to clarify something. I don't want to just search the web, I could have done that, so I would have liked a question-answer to be

more of a dialogue. So Siri would ask me a question, what kind of, "are you making this now or would you like me to create a shopping list for the ingredients?" [FG3, P2]

P3:"Yeah. And if they can chat like a human being, it would be more fun, in that state." P2:"I wouldn't. It's that like, they could talk whenever they wanted?" P1: "Sometimes I just want to give it a command, and sometimes I'll make it more of a conversational question, because that's how I'm thinking of it" [FG5]

It is clear that our users imbue Siri with human traits such as personality. The nationality of the voice chosen seemed to drive perceptions of personality, potentially giving very different types of user experience. Although the impact of this personality was generally positive, there were strong reservations from our users as to whether humanness was the most effective metaphor for interaction, with Siri still being very limited compared to humans. Equally, the future prospect of Siri being able to do more than a human received mixed views.

Trust, Data Privacy, Transparency and Data Ownership

Issues of trust regarding the use of Siri were prevalent amongst participants. Trust was discussed in the context of reliability and consistency of Siri's performance. However, we also uncovered issues of data privacy and confidentiality and significant concern around monetization, data permanency and issues of transparency.

Reliability & Consistency of Performance

Several participants reported not trusting Siri even for simple everyday tasks such as setting alarms and making calls as they cannot rely on it to perform consistently.

Especially setting reminders and alarms and things like that. Not at all. Many times Siri has called the wrong person, so things like that. (…) *I think that's part of the problem with it, is I don't think we trust it. Because if I ask it to set an alarm, I'm probably checking that alarm got set. And that probably defeats the entire purpose of its existence."* [FG3, P3]

Data Privacy & Confidentiality

Trust was more of an issue amongst participants when the task at hand involved sensitive data. What participants considered as sensitive data with respect to Siri was a surprising finding. Health and banking information were seen as sensitive, but were not ranked as sensitive as interpersonal information – such as personal contacts or announcing a significant life event – or job related information or tasks.

"*Yeah. I think there's very few things that I wouldn't trust Siri to do, but I think it wouldn't be things like, protecting my [banking] data or protecting my information. It would be more interpersonal things, or things that I valued that aren't inherently monetarily valued or something. Like I wouldn't trust Siri to send a signed copy of that contract to whoever. I would want to double-check on something that was important. Or something interpersonal, like a message to a partner or something like that, that was very important, or a birth announcement, something like that.* [FG3,P2]

With respect to Siri handling confidential – mostly financial – information, participants' views varied. Some had no issue with Siri handling such data. They felt it was no different to having that information in another app on your phone or as part of your Apple account. Some were strongly opposed to the idea of Siri handling any such information.

Monetizing, Data Permanency & Tracking

Amongst the reasons for not trusting Siri with sensitive data were significant concerns around data tracking and monetizing practices. Most of the participants were certain that "*Apple are listening to everything that's said and keeping data on everything that's said.*" [FG3, P2] and several expressed concerns around the company profiting from such data:

"*One of my biggest problem with the trust in Siri would be watching what you say to it. I mean I could be just joking around and you can say anything to it but if that's being stored and put into a database and possibly sold for advertising, I mean that's unethical.* [FG1, P2]

An uncertainty around what happens to users' data and an overall lack of transparency around Siri's and Apple's operations was also a major concern in our data, expressed even from participants who had stated that they don't mind Siri having access to confidential information:

"*But they don't talk about it, so we don't know and we have doubt, whether they will store it or not, or keep it or not. They will build up the algorithm like she was saying, we don't know, so that's why we have doubt.*" [FG5, P1]

"*….because I don't even know whether the stuff I say to Siri is being collected I mean I assume it's going to be collected by the virtue that it was signed off you know the user agreement when I got the phone"* [FG1, P3]

Several participants admitted to "*turning everything off"* after that they found out *"that Apple is tracking your recording and stuff like that"* [FG1, P3] as a way to avoid being tracked. A few participants suggested ways that this could be addressed, such as Apple or similar companies commiting publicly to not use the information for marketing and to espouse transparent practices. At the same time they admitted to feeling helpless in that to expect such changes would be an "*illusion*":

"*I don't think you'd be under the illusion anymore that whatever the hell you're doing on the Internet or any technology isn't recorded by someone somewhere. So I think you just kind of give it up, you know even when you open an incognito browser now and it's like "no, remember your boss can see this." So I think we've just kind of given up on that. You either do stuff with technology or you're completely anti it."*[FG3, P3]

Based on our participant comments, fears over privacy and data permanency are strong reasons why people may not be engaging further with IPAs like Siri. This clearly needs to be carefully addressed in relation to the IPA user experience.

DISCUSSION

Our study highlights a number of reasons why infrequent users do not engage more fully with IPAs such as Siri. Interrupting hands free interaction was seen to negatively affect how useful these users thought Siri was. They were also frustrated at the lack of customization and integration with third party apps and services that they commonly use. Our data also emphasized the importance of social embarrassment and cultural context in stopping these users from using IPAs in public spaces, limiting the usefulness of IPAs as a mobile application. Concerns over data privacy, ownership and use of user data were also strongly emphasized. The human like nature of IPAs like Siri improved the experience to some extent yet our users also saw them as limited in their humanness, raising questions as to the need and desire to use humanness as an interaction metaphor. These findings and the implications they have for the development of IPAs are discussed in more detail below.

Interruptability and Integration

Our study found that for infrequent users, hands free interaction was the major use case and in this context interruptability was a significant barrier. Our participants were frustrated when they were asked by the IPA to engage visually with the screen, or to confirm or select options by taping the touchscreen rather than through using speech. This seemed to significantly interrupt the hands-free experience of IPAs and was considered particularly problematic in situations such as driving. We suggest that maintaining speech as the main input and output throughout interaction needs to be a priority in future design of IPAs in order to ensure that hands free interaction is supported fully and that tasks are not interrupted by an interaction modality shift.

Previous work echoes our findings, highlighting the negative effects of using visual menus with benefits of spoken menu prompts in speech interface interaction more generally [21]. Based on this, and our findings, the use of speech and the minimsation or complete removal of screen based interaction should be considered especially in contexts where hands free operation is required. For instance, using an element of context awareness (e.g. identfying when the app is being used in a car or when there is likelihood to be a high attention primary task), the IPA could switch to a speech only mode, rather than requiring screen input, so that the interaction becomes more suited to the user's situation.

Our findings also suggest that improving interruptability should equally be considered in the context of IPAs' integration with other apps, platforms and devices. Users were frustrated at a perceived lack of integration with third party apps, specifically the apps that they use on a regular basis, without clear options to choose or customise which apps were used as defaults. When they could, the way to use other apps with Siri was also seen as unnatural and opaque. Typically IPAs are depicted as being agents acting on users' behalf, as an intermediary, taking care of requests by using different applications/scenarios. The current reality of Siri in particular falls short of this. While Apple has recently made attempts to address this [43] by allowing more interaction between their IPA and other apps, our findings suggest that there is a broader conversation that needs to made around issues of control, ownership and customisation of IPAs. This type of proprietary control of the channel of the information and services has been the topic of recent work in HCI [2]. Limiting access to only proprietary services and apps was seen as a significant barrier to our users in the usefulness of IPAs. Therefore for IPAs to become truly useful to infrequent users, further work needs to be done to integrate IPAs with the applications and services that they use frequently.

Social embarrassment as a barrier to using IPAs

Our study also showed that infrequent users did not feel comfortable using IPAs like Siri in public, limiting their use as a mobile application. Prominent examples of where our participants felt they would not use IPAs were on buses, offices or in business meetings, all highly public spaces. Other work has mentioned that social contextual issues could be an important consideration in speech interface [40] and IPA use [33,34]. Our work confirms that and further sheds light on why infrequent users feel that IPA use is not acceptable in a social setting. We found that they were very concerned about social judgement and embarrassment. Our users worried that using Siri in the locations mentioned would be awkward or socially embarrassing and would be against the social norms of their particular cultural background. In addition to this, they also had a clear concern around social interruptability. By using IPAs in public spaces or having to leave a space to use an IPA (as some participants mentioned), these users were concerned about annoying or interrupting others that were present in the same space. These findings echo issues around people's self-consciousness and interactions in public spaces, by Goffman [19] and equally in Humphreys' [22] ethnography on cellphone use in public spaces.

While such issues have been reported in the past with mobile phones [22], new technologies bring new social rules and dilemmas that need to be considered. The use of IPAs in public is similar in this regard. For example, given the expansion and pervasiveness of speech assistants nowadays, we need to consider ways to mediate issues of social embarrassment and the sharing of socially sensitive information. Equally, like norms around cellphone use in public vary across cultures [23,26], there may be cultural variability around norms for IPAs in these contexts. As our work highlighted, a number of participants commented on

the fact that they would not use Siri in public because of their *[specific nationality]ness*, making the cultural context a key aspect to consider in improving the IPA user experience for these users. What is clear from our findings is that social concerns are a significant obstacle to increasing IPA use for infrequent users. Although this may shift as IPA use in public becomes more socially acceptable, it is currently a significant obstacle to their wider use, potentially driving infrequent usage patterns.

Issues of trust, data permanency and ownership

Trust was also a key finding of our data analysis. Similar to frequent users [31] our participants mentioned consistency and reliability issues around Siri's ability to execute a task, especially in situations where Siri needed to use sensitive or interpersonal data. Improving the consistency of the performance of IPAs is therefore a first step towards building trust and increasing usage frequency. Similarly, considering the effect of accent on user credibility [14], improving pronunciation inaccuracies in the speech synthesis can also imbue trust.

However, unlike work on frequent users, our analysis revealed much deeper trust concerns from infrequent users over what companies who create IPAs do with the collected data. Our participants were unsure whether data was indeed being collected and stored and if so for what purpose. Many hypothesised that the data was being stored and sold to marketing organisations to monetize the interaction. Interpersonal data, rather than banking or health data were also seen as more sensitive for IPAs to have access to and handle in tasks. Such a concern needs much wider debate in the field. For instance, a number of IPAs use wake words (eg. "Hey Siri") meaning that their microphones are on constantly to catch these words when uttered. Amazon is clear in its use of previous utterance recordings to improve future responses and system performance, whilst also giving users access to recordings of their commands. Our participants' suspicion of companies and personal data use is reminiscent of issues seen in work on home energy management agents [37], where people were concerned about how companies might exploit their data, which they clearly felt belonged to them, and use it for commercial gain.

This trust barrier and suspicion needs to be addressed in the development of IPAs in the future. We suggest that IPA developers make clear statements to the user on what their data will and will not be used to do in the interaction, with particular emphasis on interpersonal data. In addition, giving users access to these recorded utterances and providing options for what their speech data are to be used for by opting in and out of specific activities can alleviate the issues of trust around this issue. Equally, as technologies evolve - new features are being added and new apps are being integrated - developers of IPAs need to engage with a more nuanced way of consent. In this respect it might be relevant to consider a continuous process of consent (instead of one-off) similar to the one suggested in Rodden et al. [37]. These may help alleviate some of the data usage and privacy concerns that seem particularly precient for infrequent users in our study.

The limitations of humanness as a metaphor

Our participants clearly understood that Siri, like most IPAs, was designed to be seen as human-like. People used this as a prism to understand the interaction, informing the users' model of what they could and could not do. They imbued Siri with intelligence and personality, with people seeing Siri as being "sassy" and "friendly". Some noted that its human-like qualities affected how they felt towards it in that they didn't want to hurt its feelings.

The synthesized voice used played a major role in influencing users' perceptions of personality. Our participants talked about a perceived shift in the personality of Siri based on the accent, mostly focusing around using different national voices. The voice, particularly the accent used, seemed to have a significant impact on their experience with the system. As mentioned earlier, accent can play an important role in affecting levels of trust for speech based output [14]. Indeed in human-human dialogue, the accent of our partner acts as a cue to our partner's knowledge state, communicative ability and social status [10]. In the same way, it is possible that the accent has a profound effect on the trust judgements and communicative attributions our infrequent users make about IPAs.

Equally, we found that humanness as a metaphor for interaction had its limitations, especially when considering the types of dialogue our users said they had with the IPA. Our participants felt like they had to "*speak like a telegram*" and alter their language choices (similar to behaviours observed in other human-computer dialogue work [4,5,27,35]), with real ambiguities as to whether Siri could understand more natural speech or could understand more complex commands that were made up of a number of turns. These behaviours are caused by a perceived mismatch between the system's capabilities and the user's model of how the system works [35,36]**.** This humanness may set unrealistic expectations [15,31]. Some of our participants actually questioned whether this metaphor was really the best way forward, seeing little need for IPAs to be like a person to achieve what users needed, echoing similar sentiments from previous research [8]. We propose that using the design of the interlocutor to signal abilities closer to the system's actual capabilities could help manage user expectations and act as a foundation for a more realistic user mental model and improved interaction. For instance, using a less human-like voice that signals more basic conversational abilities (e.g. [13]) may facilitate a mental model that is closer to the true abilities and level of task complexity conducted by the IPA, improving the user experience.

That said, there was a clear desire from some of our infrequent users for their interaction to be more conversational (a major challenge currently being addressed in the speech technology community [17,18]). IPA developers must consider whether, for the specific task that their IPA is designed, humanness is indeed the right interaction metaphor to be signalling to users as the mismatch between expectations and system performance may be affecting people's levels of IPA use.

FUTURE CHALLENGES & CONCLUSIONS

While the availability and technology that enables IPAs has advanced rapidly in recent years, research around the user experience of IPAs remains in its infancy. Our own study has its limitations such as the relatively homogeneous sample, with a number of our participants being students, although non-students were also included. Focusing on infrequent users of Siri has allowed us to identify a number of issues that these users see as significant barriers to using IPAs regularly. Of course infrequent users expressed some common issues with frequent IPA users [31] e.g. the difficulties caused by interruptions to hands free interaction, the tendency to imbue Siri with human-like qualities and the lack of trust in Siri performing tasks. The potential drawbacks of using humanness as a metaphor for interaction with IPAs were also clear. Yet these seem to be significant barriers to further use of an IPA by infrequent users, whereas frequent users seem less impacted by these issues.

Along with the findings above our work uncovers a number of unique barriers to IPA use that seem to be important to infrequent users. In particular, uncertainty around data permanence and the ownership and monetization of users' interaction data emerged as a key barrier to regular adoption for our users. These are significant challenges in the field that must be more fully understood. Specific studies looking at how people's views around privacy and trust differ in this domain and how these map to wider views of privacy in interaction could give us a deeper understanding of this important theme.

Our users also emphasised how the behaviour of IPAs influenced their speech. Too little is understood about the design of IPAs and how our experiences in these interactions causally affect our language choices in IPA interaction. Like others [31], we feel that much needs to be done to understand our mental models of IPAs and, importantly, how this affects our interaction. Our work here highlights a number of views about Siri related to accent, task performance and human-likeness that could profoundly influence people's mental model of IPAs. We suggest that these may be fruitful avenues to explore.

This paper has focused on IPA use in mobile contexts. Yet IPAs have recently become a primary interface in a number of new products, including smart home and in-car devices. These developments may lead to a different set of user concerns and barriers for user adoption or vary the prominence of barriers we found in our work. For instance, devices such as Google Home and Amazon Echo use an IPA as the primary form of communicating content and controling applications without the use of a screen. These therefore remove the barrier of having to use the touchscreen in interaction highlighted in our work. The role of context also needs considerable attention in this regard. Devices such as Amazon Echo are predominantly being placed in home spaces such as living rooms and kitchens [44], potentially amplifying concerns over social embarrassment and interrupting others compared to when users are able to leave the room easily with a mobile based IPA. What is more, using IPAs in an automotive context may lead the modality of interaction (e.g. whether interaction is soley speech based or includes screen based interaction) to be of more concern than other issues raised in our work. However, a number of barriers will cut across context and devices. Issues of trust, data permanency and ownership and the limitations of human likeness are likely to be ubiquitous across IPAs. Across all of these settings, addressing the experience of infrequent users will be valuable in helping IPA designers to understand the barriers that limit IPA adoption, which in turn will allow us to develop more effective and increasingly positive IPA interaction experiences.

ACKNOWLEDGMENTS

[Removed for anonymous review]

REFERENCES

1. Matthew P. Aylett, Per Ola Kristensson, Steve Whittaker, and Yolanda Vazquez-Alvarez. 2014. None of a CHInd: Relationship Counselling for HCI and Speech Technology. In *CHI '14 Extended Abstracts on Human Factors in Computing Systems* (CHI EA '14), 749–760. https://doi.org/10.1145/2559206.2578868
2. Matthew P. Aylett and Shaun Lawson. 2016. The Smartphone: A Lacanian Stain, A Tech Killer, and an Embodiment of Radical Individualism. In *Proceedings of the 2016 CHI Conference Extended Abstracts on Human Factors in Computing Systems* (CHI EA '16), 501–511. https://doi.org/10.1145/2851581.2892581
3. Christopher Baber. Developing interactive speech technology. *Interactive speech technology*: 1–18.
4. Linda Bell and Joakim Gustafson. 1999. Interaction with an animated agent in a spoken dialogue system. In *Proceedings of the Sixth European Conference on Speech Communication and Technology*, 1143–1146.
5. Linda Bell, Joakim Gustafson, and Mattias Heldner. 2003. Prosodic adaptation in human-computer interaction. In *Proceedings ICPhS 2003*, 2453–2456. Retrieved June 18, 2013 from http://su.diva-portal.org/smash/record.jsf?pid=diva2:458528
6. H. P. Branigan, M. J. Pickering, J. M. Pearson, J. F. McLean, and A. Brown. 2011. The role of beliefs in

lexical alignment: Evidence from dialogs with humans and computers. *Cognition* 121, 1: 41–57.

7. Virginia Braun and Victoria Clarke. 2006. Using thematic analysis in psychology. *Qualitative Research in Psychology* 3, 2: 77–101. https://doi.org/10.1191/1478088706qp063oa
8. J. Cassell. 2007. Body Language: Lessons from the Near-Human. In *Genesis Redux: Essays on the History and Philosophy of Artificial Life*. University of Chicago Press, Chicago, 346–374.
9. W. Chan, N. Jaitly, Q. Le, and O. Vinyals. 2016. Listen, attend and spell: A neural network for large vocabulary conversational speech recognition. In *2016 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, 4960–4964. https://doi.org/10.1109/ICASSP.2016.7472621
10. H. H. Clark. 1996. *Using Language*. Cambridge University Press.
11. Phil Cohen, Adam Cheyer, Eric Horvitz, Rana El Kaliouby, and Steve Whittaker. 2016. On the Future of Personal Assistants. In *Proceedings of the 2016 CHI Conference Extended Abstracts on Human Factors in Computing Systems* (CHI EA '16), 1032–1037. https://doi.org/10.1145/2851581.2886425
12. Erica Cooper, Alison Chang, Yocheved Levitan, and Julia Hirschberg. 2016. Data Selection and Adaptation for Naturalness in HMM-based Speech Synthesis. *Interspeech 2016}*: 357–361.
13. B. R. Cowan, H. P. Branigan, E. Bugis, M. Obregon, and R. Beale. 2015. Voice anthropomorphism, interlocutor modelling and alignment effects on syntactic alignment in human-computer dialogue. *International Journal of Human-Computer Studies* 83: 27–42.
14. Benjamin R. Cowan, Derek Gannon, Jenny Walsh, Justin Kinneen, Eanna O'Keefe, and Linxin Xie. 2016. Towards Understanding How Speech Output Affects Navigation System Credibility. In *Proceedings of the 2016 CHI Conference Extended Abstracts on Human Factors in Computing Systems* (CHI EA '16), 2805–2812. https://doi.org/10.1145/2851581.2892469
15. Jens Edlund, Joakim Gustafson, Mattias Heldner, and Anna Hjalmarsson. 2008. Towards human-like spoken dialogue systems. *Speech Communication* 50, 8–9: 630–645. https://doi.org/10.1016/j.specom.2008.04.002
16. Factor1. 2016. Voice Assistant Anyone? Yes please, but not in public! *Creative Strategies, Inc*. Retrieved August 26, 2016 from http://creativestrategies.com/voice-assistant-anyone-yes-please-but-not-in-public/
17. Emer Gilmartin, Francesca Bonin, Loredana Cerrato, Carl Vogel, and Nick Campbell. 2015. What's the Game and Who's Got the Ball? Genre in Spoken Interaction. In *2015 AAAI Spring Symposium Series*.
18. Emer Gilmartin and Nick Campbell. 2014. More Than Just Words: Building a Chatty Robot. In *Natural Interaction with Robots, Knowbots and Smartphones*, Joseph Mariani, Sophie Rosset, Martine Garnier-Rizet and Laurence Devillers (eds.). Springer New York, 179–185. https://doi.org/10.1007/978-1-4614-8280-2_16
19. Erving Goffman. 1971. *Relations in public: microstudies of the public order*. Basic Books.
20. Johann Hauswald, Michael A. Laurenzano, Yunqi Zhang, Cheng Li, Austin Rovinski, Arjun Khurana, Ronald G. Dreslinski, Trevor Mudge, Vinicius Petrucci, Lingjia Tang, and Jason Mars. 2015. Sirius: An Open End-to-End Voice and Vision Personal Assistant and Its Implications for Future Warehouse Scale Computers. In *Proceedings of the Twentieth International Conference on Architectural Support for Programming Languages and Operating Systems* (ASPLOS '15), 223–238. https://doi.org/10.1145/2694344.2694347
21. K.S. Hone and C. Baber. 2001. Designing habitable dialogues for speech-based interaction with computers. *International Journal of Human-Computer Studies* 54, 4: 637–662. https://doi.org/10.1006/ijhc.2000.0456
22. Lee Humphreys. 2005. Cellphones in public: social interactions in a wireless era. *New Media & Society* 7, 6: 810–833. https://doi.org/10.1177/1461444805058164
23. Mizuko Ito. 2005. Mobile Phones, Japanese Youth, and the Re-placement of Social Contact. In *Mobile Communications*. Springer London, 131–148. https://doi.org/10.1007/1-84628-248-9_9
24. Kristiina Jokinen and Michael McTear. 2009. Spoken Dialogue Systems. *Synthesis Lectures on Human Language Technologies* 2, 1: 1–151. https://doi.org/10.2200/S00204ED1V01Y200910HLT005
25. Kamitis. 2016. *Itelligent Personal Assistant- Products, Technologies and Market: 2017-2022*.
26. James E. Katz and Mark Aakhus. 2002. *Perpetual Contact: Mobile Communication, Private Talk, Public Performance*. Cambridge University Press.
27. Alan Kennedy, Alan Wilkes, Leona Elder, and Wayne S. Murray. 1988. Dialogue with machines. *Cognition* 30, 1: 37–72. https://doi.org/10.1016/0010-0277(88)90003-0
28. J. Kitzinger. 1995. Qualitative research. Introducing focus groups. *BMJ (Clinical research ed.)* 311, 7000: 299–302.
29. Pierre Lison and Raveesh Meena. 2014. Spoken Dialogue Systems: The New Frontier in Human-computer Interaction. *XRDS* 21, 1: 46–51. https://doi.org/10.1145/2659891
30. Silvia Lovato and Anne Marie Piper. 2015. "Siri, is This You?": Understanding Young Children's Interactions with Voice Input Systems. In *Proceedings of the 14th International Conference on Interaction Design and Children* (IDC '15), 335–338. https://doi.org/10.1145/2771839.2771910

31. Ewa Luger and Abigail Sellen. 2016. “Like Having a Really Bad PA”: The Gulf Between User Expectation and Experience of Conversational Agents. In *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems* (CHI ’16), 5286–5297. https://doi.org/10.1145/2858036.2858288
32. P. Milhorat, S. Schlögl, G. Chollet, J. Boudy, A. Esposito, and G. Pelosi. 2014. Building the next generation of personal digital Assistants. In *2014 1st International Conference on Advanced Technologies for Signal and Image Processing (ATSIP)*, 458–463. https://doi.org/10.1109/ATSIP.2014.6834655
33. Aarthi Easwara Moorthy and Kim-Phuong L. Vu. 2014. Voice Activated Personal Assistant: Acceptability of Use in the Public Space. In *Human Interface and the Management of Information. Information and Knowledge in Applications and Services*, Sakae Yamamoto (ed.). Springer International Publishing, 324–334. https://doi.org/10.1007/978-3-319-07863-2_32
34. Aarthi Easwara Moorthy and Kim-Phuong L. Vu. 2015. Privacy Concerns for Use of Voice Activated Personal Assistant in the Public Space. *International Journal of Human–Computer Interaction* 31, 4: 307–335. https://doi.org/10.1080/10447318.2014.986642
35. S Oviatt, J Bernard, and G A Levow. 1998. Linguistic adaptations during spoken and multimodal error resolution. *Language and speech* 41 (Pt 3-4): 419–442.
36. Sharon Oviatt, Margaret MacEachern, and Gina-Anne Levow. 1998. Predicting hyperarticulate speech during human-computer error resolution. *Speech Communication* 24, 2: 87–110. https://doi.org/10.1016/S0167-6393(98)00005-3
37. Tom A. Rodden, Joel E. Fischer, Nadia Pantidi, Khaled Bachour, and Stuart Moran. 2013. At Home with Agents: Exploring Attitudes Towards Future Smart Energy Infrastructures. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (CHI ’13), 1173–1182. https://doi.org/10.1145/2470654.2466152
38. M. Russell, S. D’Arcy, and L. Qun. 2007. The Effects of Bandwidth Reduction on Human and Computer Recognition of Children’s Speech. *IEEE Signal Processing Letters* 14, 12: 1044–1046. https://doi.org/10.1109/LSP.2007.906213
39. Martin J. Russell and Shona D’Arcy. 2007. Challenges for computer recognition of children’s speech. In *SLaTE*, 108–111.
40. Ben Shneiderman. 2000. The limits of speech recognition. *Communications of the ACM* 43, 9: 63–65.
41. Linda Wulf, Markus Garschall, Julia Himmelsbach, and Manfred Tscheligi. 2014. Hands Free - Care Free: Elderly People Taking Advantage of Speech-only Interaction. In *Proceedings of the 8th Nordic Conference on Human-Computer Interaction: Fun, Fast, Foundational* (NordiCHI ’14), 203–206. https://doi.org/10.1145/2639189.2639251
42. 98% of iPhone users have tried Siri, but most don’t use it regularly - Business Insider. Retrieved September 7, 2016 from http://uk.businessinsider.com/98-of-iphone-users-have-tried-siri-but-most-dont-use-it-regularly-2016-6?r=US&IR=T
43. Here’s How Siri Will Work In MacOS Sierra | Digital Trends. Retrieved September 20, 2016 from http://www.digitaltrends.com/computing/siri-macos-sierra-features-interface/
44. The Voice UI has Gone Mainstream | Tech.pinions - Perspective, Insight, Analysis. Retrieved February 2, 2017 from https://techpinions.com/the-voice-ui-has-gone-mainstream/46148